ECE371 Neural Networks and Deep Learning Assignment 1

Yang Zhang 22308250
School of Electronics and Communication Engineering
Sun Yat-sen University, Shenzhen Campus
zhangy2585@mail2.sysu.edu.cn

Abstract: In this assignment, we present the implementation and evaluation of a deep learning model for flower classification using a convolutional neural network (**CNN**). The model was built on the **ResNet18** architecture, pretrained on *ImageNet*, and fine-tuned using a labeled flower dataset consisting of five categories. Various data augmentation techniques were applied to enhance generalization, and a learning rate scheduler was used to improve training efficiency. The model achieved a best validation accuracy of **92.66%**, demonstrating the effectiveness of transfer learning in visual recognition tasks. Through this project, we gained deeper insights into model fine-tuning, the impact of data preprocessing, and best practices in training robust image classifiers.

Keywords: Image Classification, Convolutional Neural Networks, Data Augmentation, ResNet18

1 Introduction

The task of image classification plays a critical role in various computer vision applications, including medical imaging, autonomous vehicles, and natural scene understanding. In this project, we aim to solve a multi-class image classification problem using a dataset of flower images. Our objective is to develop a robust model capable of accurately distinguishing between five categories of flowers: daisy, dandelion, rose, sunflower, and tulip.

To address this problem, we fine-tuned a pre-trained ResNet18 convolutional neural network using the PyTorch deep learning framework. Through strategic data augmentation, transfer learning, and model optimization, our approach achieved a validation accuracy of 92.66%, demonstrating the effectiveness of our method in learning discriminative visual features from limited data.

2 Related Work

Image classification has witnessed significant progress over the past decade, largely driven by advancements in deep learning, particularly Convolutional Neural Networks (CNNs). Early architectures such as LeNet and AlexNet laid the foundation for learning hierarchical representations directly from images. The introduction of deeper networks like VGGNet and GoogLeNet further improved classification performance by increasing the depth and complexity of models.

ResNet, proposed by He et al. (2015), addressed the vanishing gradient problem in deep networks by introducing residual connections, enabling the training of networks with over 100 layers. This breakthrough has influenced a wide array of subsequent architectures and remains a standard backbone for many vision tasks.

Transfer learning, where models pre-trained on large datasets such as ImageNet are adapted to specific tasks, has become a widely adopted strategy when working with smaller, domain-specific datasets. In this project, we leverage transfer learning with ResNet18 to effectively classify flower images without training a model from scratch.

3 Method

We used the publicly available flower image dataset, structured into training and validation folders, with subfolders for each class. Each image was labeled according to its folder name, and accompanying **.txt** files provided a mapping of image paths to class indices.

Our implementation pipeline consists of the following components:

- **Data Augmentation and Normalization**: To enhance model generalization and mitigate overfitting, we applied a series of transformations: random resized cropping, horizontal and vertical flipping, rotation, color jittering, followed by normalization using ImageNet statistics.
- **Model Architecture**: We employed a pre-trained ResNet18 model, modifying its final fully connected layer to match the number of flower classes (five).
- **Training Configuration**: The training utilized stochastic gradient descent with momentum (SGD, learning rate = 0.001, momentum = 0.9) and a learning rate scheduler that reduced the learning rate by a factor of 0.1 every 7 epochs.
- **Evaluation and Saving**: We trained the model for 25 epochs, evaluating it after each epoch on the validation set. The model achieving the highest validation accuracy was saved for final evaluation.

4 Experiments

We conducted extensive experiments using the described setup, training the model on a dataset comprising five classes. The following data augmentation strategy was applied to improve robustness:

- RandomResizedCrop(224)
- RandomHorizontalFlip()
- RandomVerticalFlip()
- RandomRotation(20 degrees)
- ColorJitter()

Training Configuration Summary:

Table 1: Experimental Hyperparameters

Parameter	Value
Model Architecture	ResNet-18 (pretrained on ImageNet)
Optimizer	SGD (Stochastic Gradient Descent)
Learning Rate Schedule	StepLR (step_size = 7, gamma = 0.1)
Initial Learning Rate	0.001
Momentum	0.9
Batch Size	32
Loss Function	Cross Entropy Loss
Data Augmentation	Random crop, rotation, flip, color jitter
Input Image Size	224×224
Epochs	25
Training-Validation Split	Predefined (using folders and .txt files)

Results:

The training was conducted for 25 epochs using a ResNet-18 model pretrained on ImageNet and fine-tuned on the flower classification dataset. Table 1 summarizes the training and validation accuracy across the epochs. Initially, the model exhibited modest performance (train accuracy: 66.17%, val accuracy: 83.22%), but it rapidly improved within the first five epochs. The validation accuracy surpassed 90% by epoch 5 and peaked at 92.66% in epoch 16. This demonstrates the effectiveness of the chosen data augmentation strategies and learning rate scheduling.

Table 2: Accuracy across Training and Validation Sets

Epoch	Train Accuracy	Validation Accuracy
1	66.17%	83.22%
5	88.66%	90.56%
10	90.25%	91.26%
16	89.85%	**92.66%**
25	91.48%	90.91%

A graphical representation of the accuracy and loss trends is shown in Figure 1 and Figure 2. These plots illustrate stable convergence without overfitting.

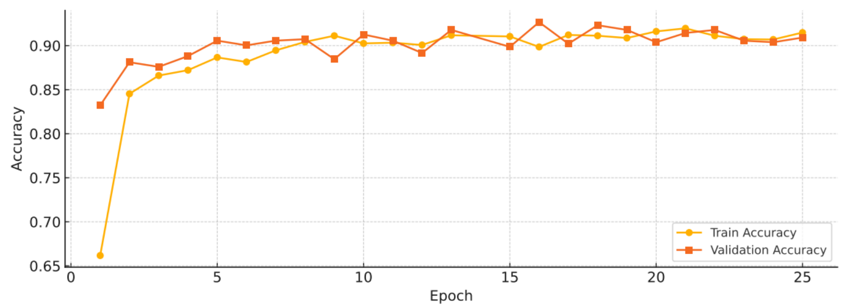

Figure 1: Accuracy Trends over Epochs

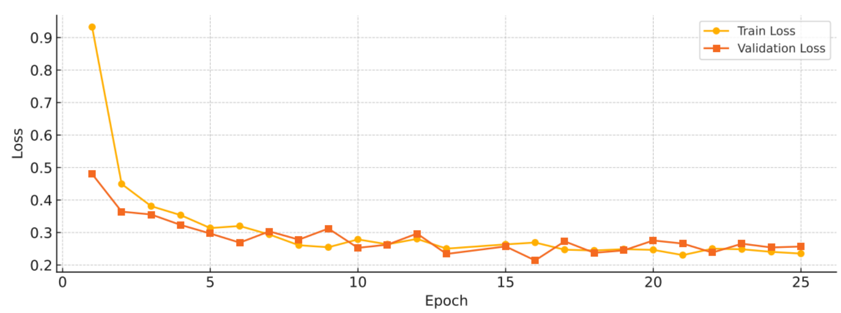

Figure 2: Loss Trends over Epochs

Discussion:

The experimental results indicate that the model quickly learns discriminative features for flower classification. The fast convergence in the early epochs can be attributed to:

- The use of a pretrained ResNet-18 model, which already possesses general feature extraction capabilities.
- Data augmentation techniques (e.g., flipping, rotation, jittering), which introduced diversity and helped the model generalize better.

The learning rate schedule played a crucial role in stabilizing performance: the initial high learning rate enabled quick learning, while the stepwise decay allowed for fine-tuning. However, there is a noticeable plateauing and slight fluctuation in validation accuracy after epoch 16, which may result from either the model reaching its capacity or the dataset's inherent difficulty.

A minor discrepancy between training and validation accuracy ($\tilde{1}$–2%) suggests that the model is not significantly overfitting, and generalization is well-controlled. Yet, further improvement may require either a deeper model (e.g., ResNet-50) or techniques such as label smoothing and test-time augmentation.

Conclusion:

This study successfully applied a transfer learning approach using ResNet-18 to the task of flower classification. The model achieved a peak validation accuracy of **92.66%**, demonstrating strong performance on a five-class classification problem with relatively limited data. Key contributing factors included effective data preprocessing, robust augmentation, and careful tuning of the learning rate.

Future work may explore additional architectural variants, such as EfficientNet or Vision Transformers, and further optimization strategies like hyperparameter search or ensemble methods. Nonetheless, the current results show that even a relatively shallow CNN, when combined with transfer learning and proper regularization, can yield competitive results on complex image classification tasks.

References